SRI International

# A GENERAL APPROACH TO MACHINE PERCEPTION OF LINEAR STRUCTURE IN IMAGED DATA

Technical Note 276

February 1983

By: Martin A. Fischler and Helen C. Wolf

Artificial Intelligence Center
Computer Science and Technology Division

SRI International®

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 859-6200 • TWX: 910-373-2046 • Telex: 334 486

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| **FEB 1983** | | **00-02-1983 to 00-02-1983** |

| 4. TITLE AND SUBTITLE | 5a. CONTRACT NUMBER |
|---|---|
| **A General Approach to Machine Perception of Linear Structure in Imaged Data** | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER |
| | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **SRI International,333 Ravenswood Avenue,Menlo Park,CA,94025** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | |
| **unclassified** | **unclassified** | **unclassified** | | **29** | |

## ABSTRACT

In this paper we address a basic problem in machine perception: the tracing of "line-like" structures appearing in an image. It is shown that this problem can profitably be viewed as the process of finding skeletons in a gray scale image after observing (1) that line detection does not necessarily depend on gradient information, but rather is approachable from the standpoint of measuring total intensity variation, and (2) that smoothing the original image produces an approximate distance transform. An effective technique for extracting the delineating skeletons from an image is presented, and examples of this approach using aerial, industrial, and radiographic imagery are shown.


## I INTRODUCTION

For many tasks in scene analysis, there may not exist general solutions independent of purpose or intended application. However, for the task of linear delineation, one can easily find image subsets for which a panel of human observers would be almost unanimous in their interpretation without having to agree on the explicit criteria underlying their decisions; our goal is to produce a computer system that can perform the delineation task at close to human levels for at least these more obvious cases, especially where semantic knowledge is not required. In this paper we present some new ways of looking at the problem of linear delineation and provide techniques that are significantly more general and effective than previously reported methods for this task.

## II PROBLEM DEFINITION

For the purposes of this paper, we define linear delineation as the task of generating a set of lists of points, for a given 2-D image, such that the points in each list fall sequentially along what any reasonable human observer would describe as a clearly visible "line-like" structure in the image. Practical examples of this task might be to delineate the roads, rivers, and rail lines in an aerial photograph, or to trace the paths taken by blood vessels in a radiographic angiogram, or to locate the wiring paths on a printed circuit board; however, our goal in this paper is not to look for specific real-world objects or to assign semantic labels to the detected linear structures, but rather to find what a human observer might choose as the most perceptually obvious

occurrences of such structures. We further distinguish between the problems of (1) detecting the edges or contours of extended objects, and (2) delineating those objects whose appearance is adequately represented by a central skeleton -- only the second problem is addressed.

## III LINES AND EDGES

While most approaches to linear delineation do not distinguish between lines and edges, and even use edge detection as a necessary first step in the delineation task, a critical concept advanced in this paper is the distinction between line and edge detection.

Edge detection is based on the concept of finding a discontinuity (in intensity or some other locally measurable attribute such as color or texture) between two adjacent but distinct regions in an image. However, in a digital representation of an image, a smooth surface can always be fit to the sample values of the integer raster. Thus, edge detection must be based on parameters or thresholds set by assumptions about the nature of the image. Even if edge points are marked only at those locations at which there are first derivative maxima, or zeros of the second derivative, ultimately, an arbitrary decision is made in deciding when the corresponding gradient is large enough to be called a discontinuity.

Intuitively, a "line-like" or linear structure is a (connected) region that is very long relative to its width, and has a ridge or skeleton along which the intensities change slowly and are distinguished from intensities outside the region; the width need not be constant, but any changes in width should occur in a smooth manner. To simplify the discussion, we will assume the linear structures are distinguished by their ridgepoints being brighter than the surrounding background, but any other specified attribute, which is locally detectable, would be an acceptable substitute. It is important to recognize the fact that a clearly visible line in an image may not have locally detectable edges and thus no locally measurable width, or possibly only one detectable edge, or even two edges which are significantly separated and nonparallel as might occur in a local widening of a river. It is also generally the case that linear structures have no visible internal detail that is essential to their delineation.

As a point of interest, it might be noted that the mechanisms for generating subjective edge and line illusions are quite different; subjective edges appear to require a 3-D interpretation, while subjective lines appear to be produced by adaptation phenomena.

## IV SMOOTHING, DISTANCE TRANSFORMS, AND THE GRAY SCALE SKELETON

If we can find the edges of a linear structure, we can generate a distance transform and extract a skeleton as the desired delineation (e.g., Rosenfeld [1], Fischler [2]). However, as noted in the preceding section, linear structures do not always have locally detectable edges, and, since all of the generally known techniques for deriving a skeleton require a complete contour, some other approach is required. The classical skeletonizing techniques intimately link the contour/edges of a region and its skeleton, and it is just this linkage that we wish to break.

Surprisingly, something equivalent to a distance transform that works on gray scale images, and on binary images as well, is already available. To achieve our purpose, we need only observe that the intensities in a properly smoothed image can be considered to be the values of an approximate distance transform. What is the best smoothing function for general use? Actually, it doesn't seem to make much difference in many cases. Most digital images have been processed by low-pass optical and electronic systems that have inserted the required minimum level of smoothing. The viewpoint that the smoothed image can be considered to be a distance transform is the essential element. However, if we start with a binary image, or a very noisy image, then additional smoothing is often desirable. Since we are not concerned with blurring edges, and we would like to eliminate or blur any structure or texture internal to the linear regions, we want the smoothing function to have a width of approximately that of the region to be delineated. If the width of the smoothing function is increased further, the thinner linear structures are eventually eliminated. Thus, if we wish to find all possible linear structures without prior knowledge of the content of the image, the processing should be repeated with a set of smoothing filters having a spectrum of widths. Actually, no more than two or three filtering steps should ever be required. For example, to trace all the linear structures (diameters up to 20 pixels) in a noisy radiographic angiogram, a single filter of width 20 was used (see Figure 1). Smoothing introduced by the acquisition process was sufficient to produce excellent results in tracing the linear structures in aerial imagery (see Figures 4 and 5).

## V RIDGES (OR VALLEYS), OPERATORS, AND NEIGHBORHOODS

Having produced an approximate distance transform via smoothing, we now must deal with the problem of locating the ridgepoints that denote the spines (skeletons) of the linear structures. When an exact distance

transform is derived from a complete contour, noise is not a problem and the skeleton has assured geometric properties that make it easy to detect; finding the ridgepoints of an approximate distance transform is considerably more complex.

We traditionally distinguish between locally and globally detectable features: local features are detectable by an intensity pattern which can be observed through a small peep-hole centered on the feature, while global features are ambiguous in a small area. The model or description of the local feature is generally compiled into an intensity patch (matched filter or operator) which can be convolved with the image to detect the corresponding feature. In the case of an exact distance transform, a 3X3 pixel operator is sufficient to detect ridgepoints (a 2X2 operator is sufficient for the Labeled Distance Transform (Fischler [2])); for the approximate distance transform, a small fixed-size operator is ineffective.

The principal utility of a local operator is that the number of data patterns the operator might encounter is small enough to allow one to enumerate a decision for each such pattern. If we further agree to use a small square window of the image as our local domain, and to use either table look-up or convolution as the basis for decision making, then a uniform mechanization can be employed to implement a large number of distinct (and generally unrelated) local operators. The attractiveness of this second implementation aspect has led to the situation that almost all low level (local) scene analysis is done using such peep-hole type operators. The disadvantage of this approach is that the concept of local is relative to the size of the entity of interest, and either one must know this size in advance, or use a whole family of operators of increasing size, where the larger operators lose the advantages that led to their use in the first place. In the case of line detection, where the line width can vary over a wide range of values, the conventional operator concept is inappropriate.

Based on these general issues (even more than on the immediate problem at hand), we have considered other realizations of general "local" decision-making processes that satisfy the previously stated conditions, but do not necessarily lend themselves to a convolution type mechanization. In particular, restricting our attention to finding the maxima and minima of functions of the displacement along a space curve defined over the image, satisfies our requirements for computational and decision-making simplicity even when the curve traverses the entire image. While the space curve might assume any shape (e.g., follow the contour of an object), the analysis itself is independent of the shape; for the linear delineation problem, we used image intensity as a function of displacement along horizontal and vertical scan lines. Since maxima and minima are symmetrical attributes, we will only discuss the problem of labeling maximal points along the curve.

The problem of finding the ridgepoints of an approximate distance transform can be viewed as the problem of finding the ridgepoints (local maxima) of an exact distance transform to which some amount of noise has been added. We are not concerned about the possibility of making isolated (incoherent) incorrect decisions, because we have developed effective linking and pruning methods, described in the next section, that are capable of eliminating such errors. Our main problem is that we cannot count on finding either large local gradients or using known line width to determine some minimum significant gradient threshold to identify valid ridgepoints; additionally, noise will introduce many false local maxima. Thus, we must use total intensity change, rather than rate of change, to detect valid ridgepoints, and we must have an effective way of determining such total change even in the presence of local variation introduced by noise. (While it is not immediately obvious that total intensity change, rather than rate of change, will recover the perceptually obvious linear features, our experiments indicate that this is indeed the case.)

Our approach is to evaluate two attributes of each of the detected intensity maxima along the space curves (in this case, horizontal and vertical scan lines), which we call the "local" and "global" maxima values. The local maxima value of a point is the total intensity difference from the point to the highest of its immediate left and right intensity minima along the curve. The left (right) global maxima value of a point is the total intensity difference from the point to the lowest intensity value found moving to the left (right) prior to encountering a point with an intensity value greater than that of the given point; the global maxima value of the point is the smaller of its left and right global values. In the case of a plateau, only the center point is treated as a maximal point and evaluated as previously described. If a point (or points) on a plateau has an immediate neighbor with a higher intensity value, it is not a maximal point and it is not assigned either a local or global value (actually, for implementation purposes, non-maximal points are assigned zero values); on the other hand, every maximal point will have both a local and global value where the global value equals or exceeds the local value. Figure 2 provides some examples illustrating the operations just described.

We have been proceeding under the assumption that a large local intensity maxima (LIM) denotes a significant event, but, in the presence of large variations in image intensity or noise, the global intensity maxima (GIM) would be a better detector of significant intensity variation; however, in a well-smoothed or relatively noise-free image, there might be very little difference in the information contained in the LIM and GIM measures. There is also the issue of deciding what is a large-enough value, of either the LIM or GIM, to indicate significance. In our unsmoothed image data, about 1/3 of the points were maximal points, and, in a smoothed image, this percentage is much smaller (see Table 1). Given the linking and pruning techniques we describe in the

next section, it might be possible to return all the maximal points in a binary mask and still extract the desired line structure from the background noise contained in such a mask. However, it makes much more sense to first eliminate those maximal points that do not have enough intensity variation to be perceptually distinguishable from a flat background. (It would even appear that we could, without losing essential information, eliminate those maximal points with a total variation less than that required to perceive them against a random-noise field with the same statistical variation as the measured variation over some surrounding neighborhood in the image.)

Rather than attempting to find some optimal threshold setting (in the sense of maximum noise elimination without any loss of linear structure), which would be difficult or impossible to automatically determine at this level of information organization, we iteratively adjust our threshold settings to satisfy a constraint based on a complexity measure. These program-determined thresholds typically allow at least two to ten times the number of ridgepoints (maxima) to be retained above that which would result by manually setting the thresholds to achieve visually acceptable results. The final elimination of "non-significant" maxima is achieved later in the processing at a higher level of organization.

## VI CLUSTERING, LINKING, PRUNING, NODE ANALYSIS, RANKING, AND FINAL DELINEATION

Based on the availability of a binary overlay depicting the locations of the major linear structures contained in the given gray scale image, obtained as described in the preceding section, we have been able to demonstrate that.

(1) The linking step in the delineation process can effectively be based on the single attribute of geometric proximity, and that a clustering or association step, followed by the construction of a Minimum Spanning Tree (MST) through the points of each cluster (Fischler [3], Zahn [4]), will correctly link the ridgepoints along the skeletons of the linear structures.

(2) The desired delineations will be embedded in trees containing additional branches that are either minor linear structures or noise, and that simple pruning techniques can eliminate most of this unwanted detail (see Figure 3; note that tree pruning can effectively achieve simplifications that would be difficult, if not impossible, at lower levels of organization of the information).

(3) Having properly linked the ridgepoints and pruned some of the smaller branches of the resulting trees, we can extract long coherent paths by a decision procedure applied at each node of each tree. This decision procedure, based on the local branch attributes of intensity, connectivity, and directionality at each node, assigns path connectivity through a node by splitting off incompatible branches; any remaining ambiguities (more than two branches entering a node) are resolved by choosing those pairings that result in the longest paths.

(4) The paths obtained in the tree partitioning step can be rank ordered with respect to perceptual quality by a metric based on the path attributes of total length, contrast, and continuity.

Details of the procedures discussed in this section are presented in the appendix.

## VII EXPERIMENTAL RESULTS

Figures 4-10 present the results of applying the delineation algorithm to aerial, industrial, and medical images. Our goal was to be able to take imagery from arbitrary domains, and without any human intervention (e.g., parameter adjustment or attention focusing), produce high quality delineation of the obvious linear structures. The results show that we have accomplished much of what we originally intended. The delineations achieved by the uniform parameter settings are quite good in all the images with the exception of the angiogram (which is extremely noisy, and does not really satisfy the criteria of having "clearly visible" linear structure); and even here, by making an appropriate selection of two parameters, the smoothing diameter and the association distance (required in the clustering step), we obtain very good results. By using the values produced by the ranking step of the delineation algorithm, we believe it should be possible to automatically search the parameter space (of approximately 10 to 100 parameter combinations) to optimize the processing for any given image. We are currently investigating this possibility.

## VIII CONCLUDING COMMENTS

We have presented the viewpoint that the problem of delineating the obvious linear structures in an image is distinct from that of finding edges or contours, and is best viewed as the process of finding skeletons in gray scale images (i.e., that line detection does not necessarily depend on gradient information, but rather is approachable from the standpoint of detecting total intensity variation); as a necessary step in this process, we have suggested that an approximate gray scale distance transform can be attained by smoothing the original image. We have described an effective technique for finding ridgepoints (points on the delineating skeleton), and, in the process, raised some important questions about the conventional approach to designing "local operators."

The competence of the ridgepoint algorithm to abstract the linear structure of an image is apparent by inspection of Figures 4 through 10. This effectiveness is, in a sense, a "psychological discovery." It would appear that ridgepoints are important perceptual primitives which may play a significant role in a variety of other tasks (e.g., perception of surface shape). A significant feature of "ridgepoints" is that their locations are independent of any monotonic intensity transformation of the image, and their geometric configuration is independent of a change of scale. These are precisely the properties we would expect of a perceptual primitive, but are lacking in the commonly employed gradient dependent primitives, such as "edgepoints."

Starting with both the binary overlay (produced as discussed in the main body of this paper) and the original gray scale image, we have demonstrated via examples that the remaining steps in the delineation process can be effectively achieved.

Our goal in this work has been to approach human levels of performance in finding perceptually obvious delineations in images selected at random from a reasonably broad class of scene domains, and without any human intervention or prior knowledge about the image content. We believe that this goal can be achieved through extension and refinement of the techniques described in this paper.

## ACKNOWLEDGEMENT

The research reported herein was supported by the Defense Advanced Research Projects Agency under Contract No. MDA903-83-C-0027 and MDA903-79-C-0588, and by the National Science Foundation under Contract No. ECS-7917028.

## REFERENCES


1. A. Rosenfeld and J.L. Pfalz, "Sequential operations in digital picture processing," J. ACM, Vol. 13(4), pp. 471-494 (October 1966).

2. M.A. Fischler and P. Barrett, "An iconic transform for sketch completion and shape abstraction," Computer Graphics and Image Processing, Vol. 13(4), pp. 334-360 (August 1980).

3. M.A. Fischler, J.M. Tenenbaum, and H.C. Wolf, "Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique," Computer Graphics and Image Processing, Vol. 15(3), pp. 201-223 (March 1981).

4. C.T. Zahn, "Graph-theoretical methods for detecting and describing gestat clusters," IEEE Trans. on Comp., Vol. C-20(1), pp. 68-86 (January 1971).

5. O. Firschein and M.A. Fischler, "Association algorithms for digital imagery," Conference Record of the Twelfth Annual Asilomar Conference on Circuits, Systems, and Computers, IEEE Catalog No. 78CH1369-8C/CAS/CS (November 1978).

6. E.M. Reingold, et al, "Combinatorial Algorithms," Prentice-Hall (1977).

7. M.A. Fischler and R.C. Bolles, "Perceptual organization and curve partitioning," [in preparation, 1983].

## Appendix A

This appendix describes the procedures invoked by the delineation algorithm. These include :

a) Smoothing
b) Ridgepoint Detection
c) Association (grouping or clustering) and Thresholding
d) Linking
e) Pruning
f) Linear Segment Extraction
g) Linear Segment Decomposition
h) Linear Segment Evaluation

### SMOOTHING

a) Purpose: enhance connectivity, especially of "thick" lines, in the presence of noise (e.g., along noisy plateaus -- see Figure 11).

b) Algorithm: 2-D Gaussian smoothing approximated by a 1-D binomial mask iteratively convolved with the image in the horizontal and vertical directions (e.g., a mask of diameter five is the sequence <1 4 6 4 1>).

c) Parameter Space: mask diameter is the only adjustable parameter. Values of <0 3 5 10 20> would probably be adequate for any 256X256 image. In our experiments, we used values of 0 or 3 for all the images except the very noisy angiogram which required a smoothing diameter of 20 for best results.

d) Parameter Selection: since smoothing can alter connectivity, and thus topology, and since topology cannot be evaluated locally, there is a strong possibility that there is no low level test which can be used to select a best smoothing mask. In general, an ideal smoothing mask should have a diameter equal to the width of the line whose connectivity it is enhancing. If the mask diameter is smaller, it is less effective; if the mask diameter is larger, it will eventually eliminate or shift the line's location (see Figure 12). Ideally, one should separately process the image with each of the available smoothing masks and piece together the best delineations after a final evaluation. The precise control structure for such an approach is still under investigation. The alternative we are using at present is to set the mask size to the diameter of the the smallest width lines we wish to ensure are included in the delineation. In our experiments we used the following values:

D = 0 for low resolution aerial images where line widths of 1-2 pixels are common. We depend on the smoothing already introduced by the limited bandwidth of the image acquisition system.

D = 3 for all the remaining "normal" imagery.

D = 20 for the very noisy angiogram.

e) Comments: For a given line of nominally constant width, the quality of its delineation plotted as a function of smoothing mask diameter, is a parabolic function peaking where mask size equals line diameter. This is the only method we have found so far to get a good estimate of line width.

RIDGEPOINT DETECTION

a) Purpose: to locate a set of points which contain the points comprising the linear structures in an image. These points should be dense along the ridge-lines (skeletons) and relatively sparse elsewhere. In addition to location, for each point we also wish to obtain a measure of "contrast" between the point and the background.

b) Algorithm: ridgepoints (maxpoints) are detected by searching for local (intensity or other attribute) maxima along horizontal and vertical scan lines. Each ridgepoint is assigned two numerical values which are measures of its "local" and "global" contrast. These values are determined as shown in Figure 2, and the larger of each of the two values determined for the horizontal and vertical directions, are retained for each of the two attributes.

c) Parameter Space: with the possible exception of scan-line positioning, the algorithm for ridgepoint detection has no internal parameters.

d) Parameter Selection: A dense set of horizontal and vertical scan lines was used in all of our experiments. Since this selection produced excellent results, nothing else was considered.

e) Comments: The competence of the ridgepoint algorithm to abstract the linear structure of an image is apparent by inspection of Figures 4-10. A significant feature of "ridepoints" is that their locations are independent of any monotonic intensity transformation of the image, and their geometric configuration is independent of a change of scale.

ASSOCIATION AND THRESHOLDING

a) Purpose: To partition the ridgepoints, detected in the image, into independent coherent subsets at some maximum level of complexity.

b) Algorithm: Points are initially grouped into mutually exclusive clusters such that every point in a cluster is within some maximum 2-D Eucledian distance from at least one other point in the cluster. The technique used is a fast "two pass" algorithm described in (Firschein and Fischler [5]). If the number of points in a cluster is greater than some preset complexity threshold, then the cluster is thinned by eliminating those points with the lowest contrast (i.e., raising the "contrast threshold"); a separate contrast threshold is thus determined independently for each cluster in a image. If the elimination of low contrast points causes the distance condition to be violated, the cluster is decomposed into subclusters which now assume independent status, and the above procedure is iteratively invoked.

c) Parameter Space: The 256X256 image contains 65,000 pixel locations, and the total number of ridgepoints in a typical scene can vary from 10,000 to 25,000 points. An initial contrast thresholding is done to reduce the number of ridgepoints to about 10 percent of the image. At present, the global contrast threshold is always set to be twice the value selected for the local contrast threshold. Typical values obtained for local contrast range from 2 to 20 intensity units for a picture with an intensity range of 0 to 255 intensity units.

The complexity threshold (the maximum number of points in a cluster) is based on both the competence of subsequent techniques, and on computer system constraints. It is currently set at 1500 points. The contrast thresholds for a cluster are adjusted by the algorithm to satisfy the complexity threshold constraint.

The distance threshold employed by the association algorithm is based on the assumption that the clearly visible linear structures in the image will be continuous, and only a slight allowance need be made for acquisition and quantization noise. Values of 2 and 3 are the only distances considered; this allows gaps of 1 and 2 pixels, respectively.

d) Parameter Selection: The complexity threshold, with a setting of 1500, allows anywhere from two to ten times more ridgepoints to appear in a cluster than will be retained in the final delineation extracted from that cluster; it is extremely conservative, and even doubling or halving its value generally has almost no effect on the final delineation. The distance threshold is set to "2" for normal images, and to "3" for exceptionally noisy images which have been smoothed with a very large diameter filter mask. In our experiments, a distance threshold of 2 was used for all the images except the angiogram where we used a distance threshold of 3.

LINKING

a) Purpose: To connect the elements of each cluster into a graph (or tree) in which branches delineate potential linear structures in the image.

b) Algorithm: Ridgepoints are linked (connected to their nearest Eucledian distance neighbors) using a Minimum Spanning Tree (MST) algorithm (Reingold [6]).

c) Parameter Space: There are no internal parameters in the MST algorithm.

PRUNING

a) Purpose: To eliminate ridgepoints which do not fall along the major linear structures in the image.

b) Algorithm: Based on the assumption that insignificant linear segments (and accidental aligmments of noise that might produce ridgepoints) will result in short terminal branches, or produce gaps near the terminal ends of branches, we iteratively delete from the tree terminal branches with fewer than (1/3)X("SHORT") points, and those segments at the ends of terminal branches which have a gap prior to encountering at least "SHORT" contiguous points; when no additional points can be deleted by further iterations, those terminal branches with fewer than "SHORT" points are also deleted.

c) Parameter Space: The single internal parameter "SHORT" is set to the length of the shortest linear segment we would be willing to consider as significant in the final output. Without knowing anything about the image domain, this parameter is based on purely psychological considerations.

d) Parameter Selection: "SHORT" has been hardwired to a value of 15 for 256X256 images; this value gives generally good results both in eliminating noise and in retaining significant structure.

e) Comments: Eliminating noise segments at the spanning tree level of organization, as opposed to simply eliminating isolated or low contrast ridgepoints, is a major contributor to the excellent performance of the complete delineation algorithm. In an experiment with 20 randomly generated binary overlays, representing detected ridgepoints (ridgepoint density 10%), no randomly formed linear segments survived this filtering step.

LINEAR SEGMENT EXTRACTION

a) Purpose: To extract individual coherent linear segments from the MST.

b) Algorithm: It is assumed that the MST for a given cluster, after the pruning step, may contain a number of inter-twined perceptually significant linear segments. The problem is to determine at each node of the MST which branches should be linked to each other, and which branches should be split off as end-point (terminal) portions of separate linear segments. The algorithm considers each node of the tree in some arbitrary order; at each node it measures four attributes of each incoming pair of branches: the "angle" the branches form with their vertex at the given node, the difference between the average intensity grayscale values of the branches, whether both branches are terminal branches, and whether either of the branches has a gap immediately adjacent to the given node. A "mismatch" score is computed for each combinatorially selected pair of branches, and that pair with the lowest score is designated the "control pair."

If the score computed for the control pair is greater than a locally computed "compatibility threshold," the node will be deleted from the MST, and the MST thus decomposed into a number of smaller disjoint trees; none of the branches entering the given node will be connected to each other in the final delineation.

If the mismatch score computed for the control pair is less than the compatibility threshold, then the control pair, and any other segment which when paired with one of the segments in the control pair has a score almost as good as that of the control pair (within 5%) will remain connected to the given node. All other branches will be split off from the given node but will be joined at a corresponding node of a new tree and then subjected to the procedure just described (this allows more than one pair of branches, entering a given node, to be linked into a continuous path).

After all of the nodes have been subjected to the above procedure, the maximum length paths in all of the resulting trees will be extracted as independent linear segments. Branches split off from the maximum length paths form new trees which are recursively processed in the same way.

Parameter Space: This procedure has five parameters; these are the compatibility threshold and weighting factors assigned to each of the attributes in computing the incompatibility score for a pair of branches entering a given node. All attributes are computed using, at most, the eight points closest to the node along each of the entering branches.

a) Compatibility Threshold: computed for each node as 20% of the average grayscale value (AGSV) of all the branches entering the node.

b) Penalty for two branches meeting at an acute angle = (6%)X(AGSV).

c) Reward for two branches meeting at an angle of at least 145 degrees = (8%)X(AGSV).

d) Penalty for both branches being terminal branches = (5%)X(AGSV).

e) Penalty for either (or both) branches having a gap immediately adjacent to the given node = (4%)X(AGSV).

f) Penalty for difference in average gray scale value of the two branches = (actual gray scale difference).

Parameter Selection: The problem of measuring local similarity, and best geometric continuation, of a set of line segments entering a common node, is at a level of complexity requiring a rule-based system to make a determination in reasonable agreement with human perceptual judgements. The particular attributes chosen were ones which seemed relevant and could be easily measured (line-width would have been included as an attribute if we had a good way of measuring it locally; our method of measuring the "angle" between two segments will be improved in the near future). The selected values of the weighting factors were based on informal experiments, and then hard-wired into the system.

Comments: It is obvious that the selected attributes, our method of measuring them, and the assigned weighting factors, can all be improved by a more systematic study of their psychological implications; nevertheless, the currently implemented procedure seems to make very few poor decisions and does not appear subject to major improvement.

## LINEAR SEGMENT DECOMPOSITION

Purpose: To split individual linear segments that are not perceptually "homogenous."

Algorithm: The preceding procedure for linear segment extraction accomplishes this operation when two or more segments come together at a common point (node). However, it is possible that exactly two distinct segments can accidentally meet at their nominal end-points, and a special procedure may be required to check for this situation. No algorithm is currently included for this purpose. However, the linear segment extraction algorithm could be applied at every point along a given line segment to achieve this purpose if desired.

LINEAR SEGMENT EVALUATION

Purpose: To rank order a given set of linear segments with respect to perceptual quality (e.g., given two alternative potential delineations, which one is more acceptable from a perceptual standpoint; or assuming we wish to restrict our "sketch" of an image to the "best" N linear segments, how the selection is to be made).

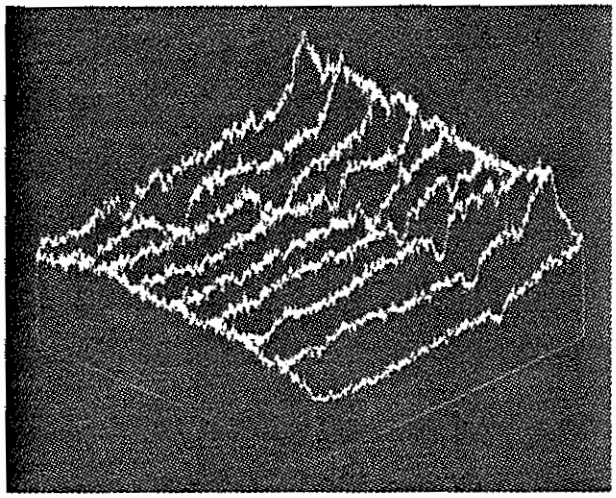
Algorithm: This procedure separately measures two attributes of each line segment -- its photometric quality (contrast with the background) and its geometric quality (coherent length and smoothness). Photometric quality is measured by simply summing the Local Maxima values assigned to each point of the segment. Geometric quality is computed by adding the number of points in the segment to the distance between end points of the segment and then subtracting three times the number of gaps, and also the number of locations at which the segment has discontinuity in local smoothness. Numerical values are normalized by converting them into rank orderings for all of the segments being evaluated. The final score assigned to each segment is the sum of its photometric and geometric orderings, with the geometric score weighted twice that of the photometric score.

Parameter Space: Four quantities are measured for each line segment: contrast, number of points, number of gaps, and number of discontinuities in smoothness. The final ranking is a weighted combination of the values of these attributes.

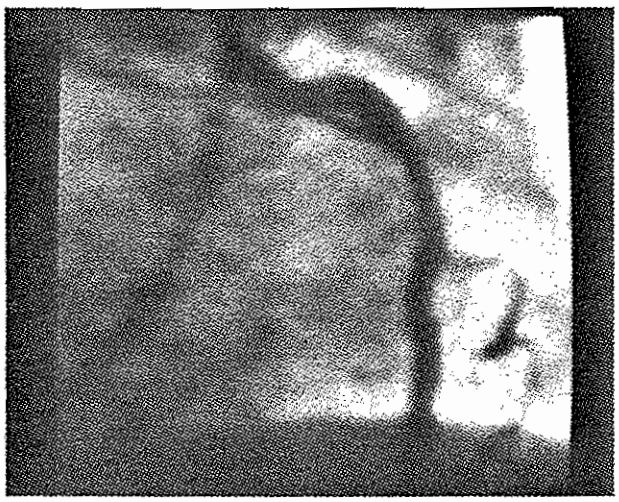
Parameter Selection: The relative weightings assigned to the attributes, as described above, were determined by informal experimentation and the hard-wired into the system.

Comments: More than any other of the preceding steps in the delineation process, evaluation must be based on psychological factors, and thus effective performance depends on observing and modeling human behavior in this task. Our presently implemented algorithm, described above, is a first unsophisticated attempt to achieve this goal -- its main defects are lack of an effective way to measure local width and local smoothness. This particular procedure, while not currently essential to the overall performance of the delineation algorithm, can play a vital role in automatically adjusting the heuristically set parameters of the other procedures through performance evaluation feedback. In order to take advantage of such an approach, the psychological model underlying this procedure, and the methods for measuring the relevant attributes, still need significant improvement. Other work now underway is addressing some of these issues (reference Fischler and Bolles [7]).

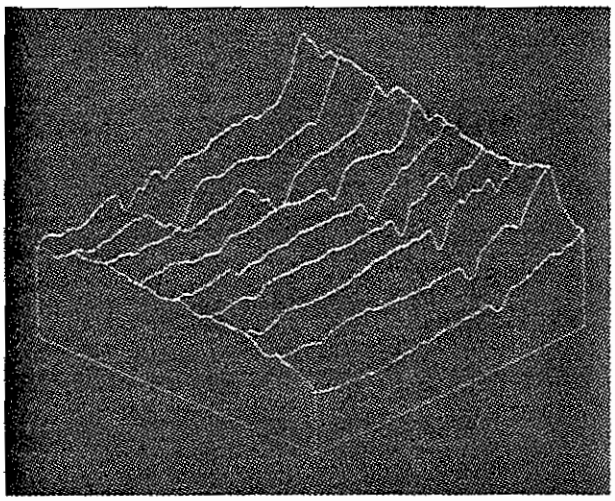
(a) ORIGINAL GRAY SCALE IMAGE



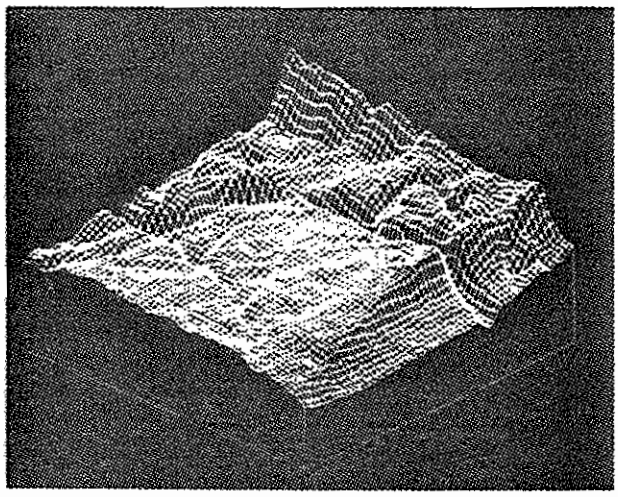
(b) INTENSITY PROFILES OF ORIGINAL IMAGE



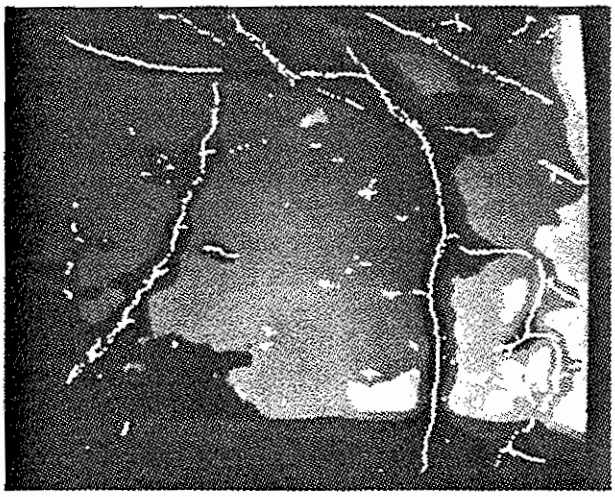
(c) SMOOTHED GRAY SCALE IMAGE



(d) INTENSITY PROFILES OF SMOOTHED IMAGE

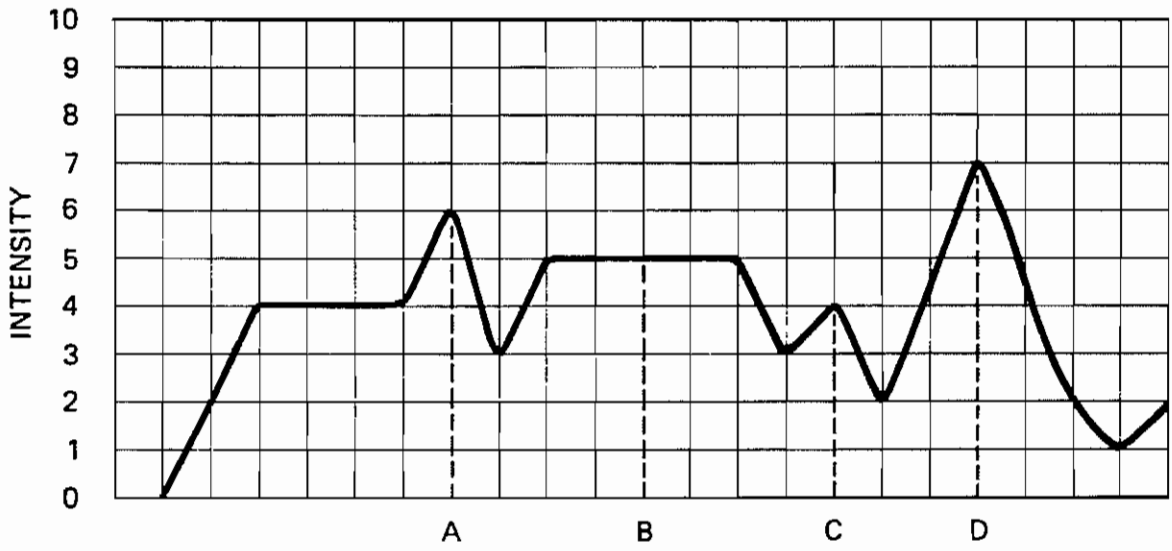


(e) TERRAIN MAP OF INTENSITY SURFACE OF SMOOTHED IMAGE



(f) LINEAR FEATURE POINTS FOUND IN VALLEYS OF INTENSITY SURFACE

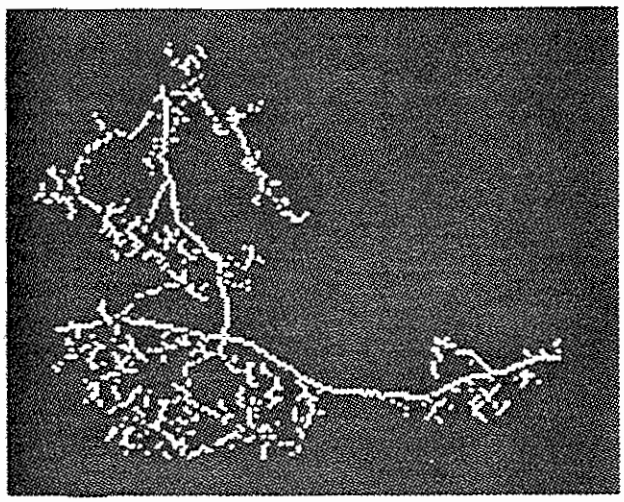
FIGURE 1 IMAGE SMOOTHING, AND FINDING LINEAR FEATURE POINTS IN THE INTENSITY VALLEYS OF A RADIOGRAPHIC IMAGE (ANGIOGRAM)



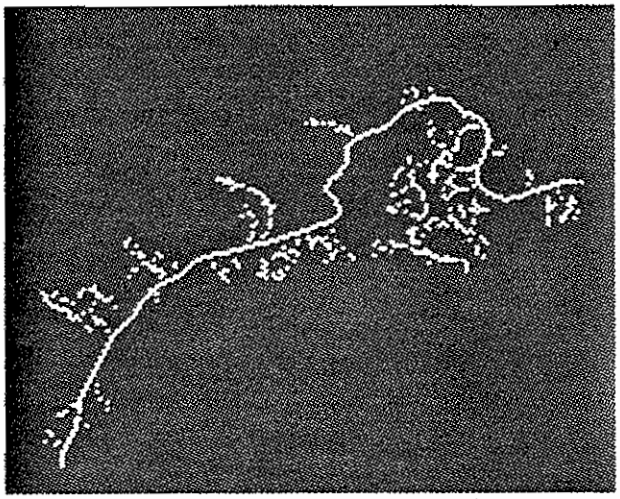

LLM = left local minimum intensity
LGM = left global minimum intensity
RLM = right local minimum intensity
RGM = right global minimum intensity
LIM = "local maxima" of intensity difference
GIM = "global maxima" of intensity difference

| Point | A | B | C | D |
|---|---|---|---|---|
| Point intensity | 6 | 5 | 4 | 7 |
| LLM | 0 | 3 | 3 | 2 |
| RLM | 3 | 3 | 2 | 1 |
| LGM | 0 | 3 | 3 | 0 |
| RGM | 2 | 2 | 2 | 1 |
| LIM | 6 - 3 = 3 | 5 - 3 = 2 | 4 - 3 = 1 | 7 - 2 = 5 |
| GIM | 6 - 2 = 4 | 5 - 3 = 2 | 4 - 3 = 1 | 7 - 1 = 6 |

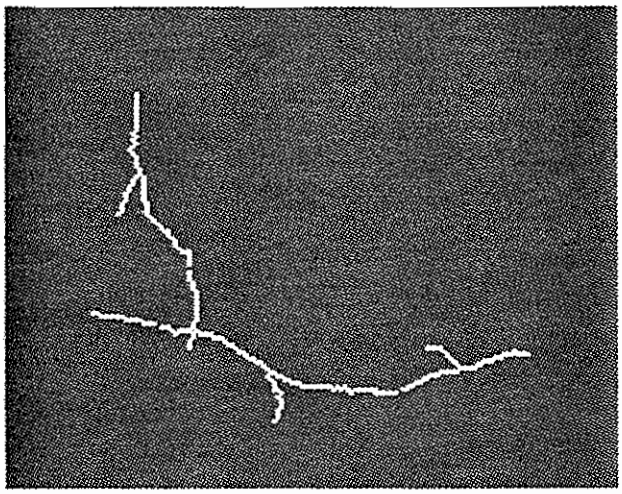
FIGURE 2 EXAMPLES OF LOCAL AND GLOBAL MAXIMA COMPUTATION



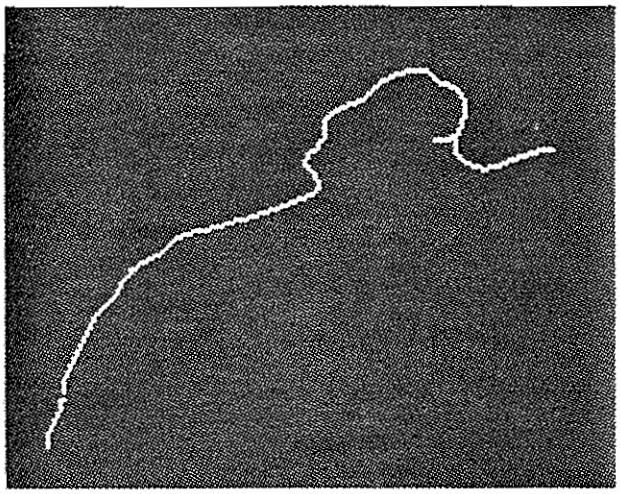
(a) A SINGLE CLUSTER OF LINEAR FEATURE POINTS



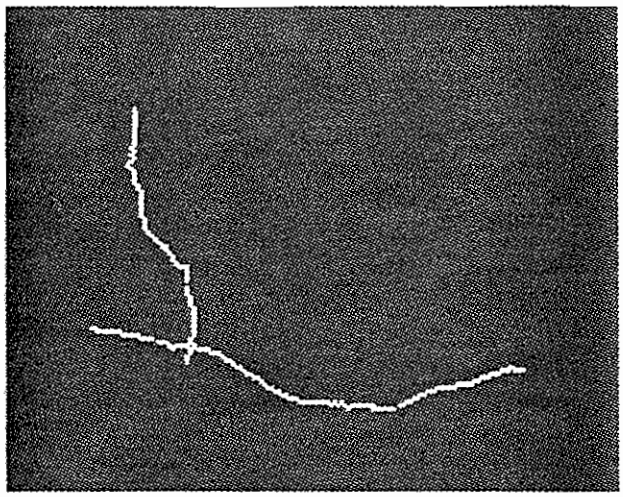
(b) A SINGLE CLUSTER OF LINEAR FEATURE POINTS



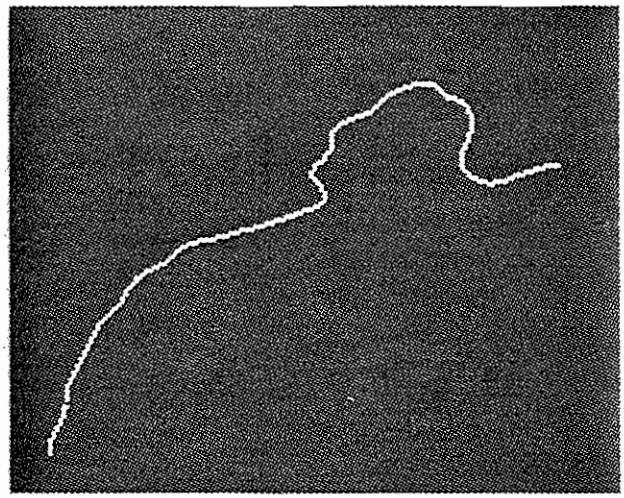
(c) CLUSTER AFTER PRUNING



(d) CLUSTER AFTER PRUNING

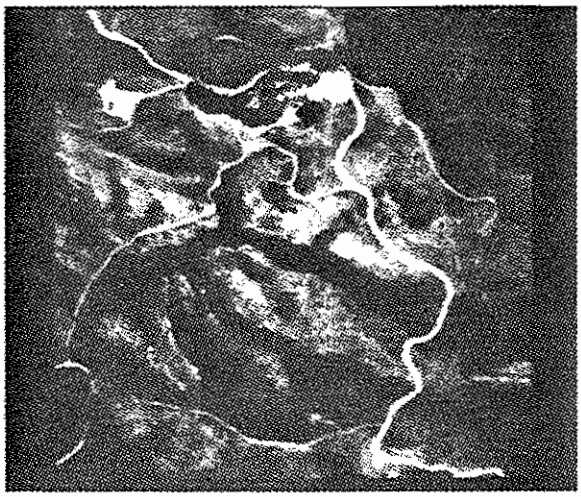


(e) THE TWO LINEAR SEGMENTS EXTRACTED
(one horizontal line and one vertical line)

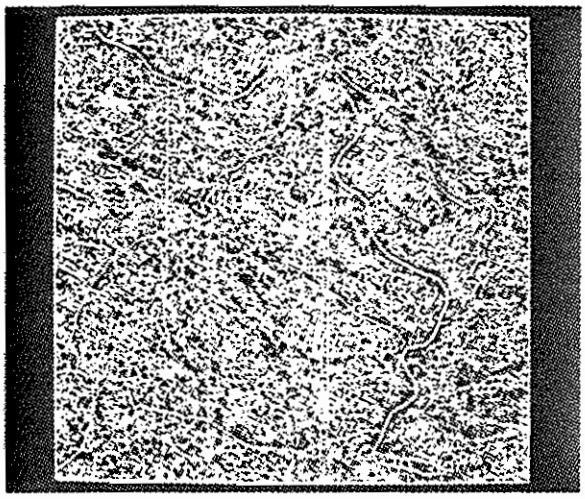


(f) THE LINEAR SEGMENT EXTRACTED

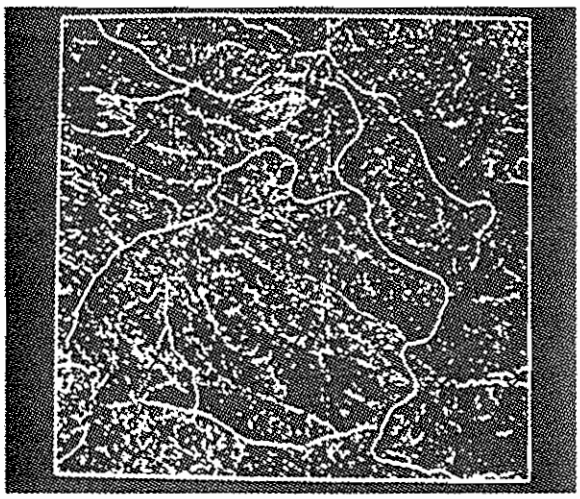
FIGURE 3 PRUNING AND LINEAR SEGMENT EXTRACTION FROM A SINGLE CLUSTER



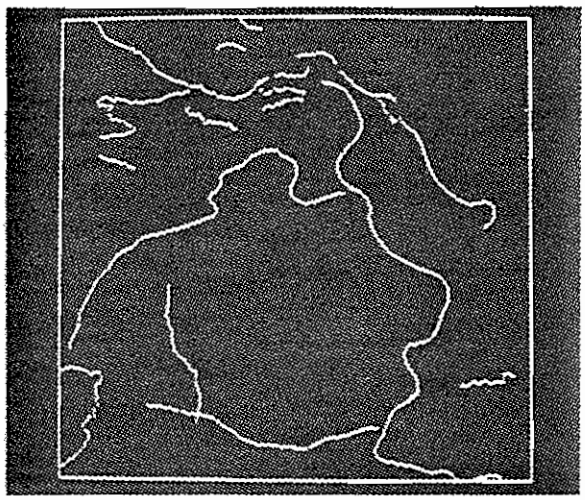
(a) ORIGINAL GRAY SCALE IMAGE: DROAD2



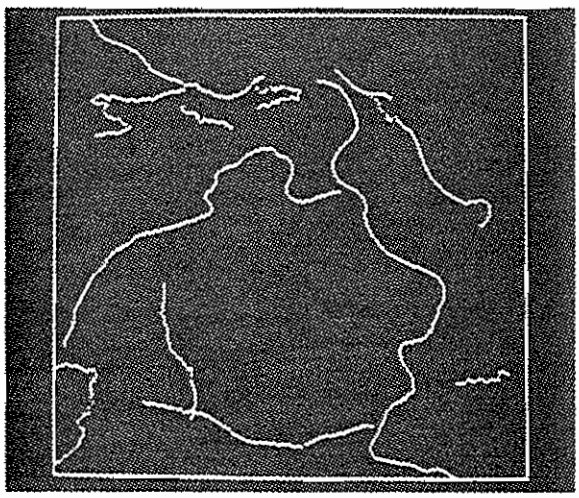
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS



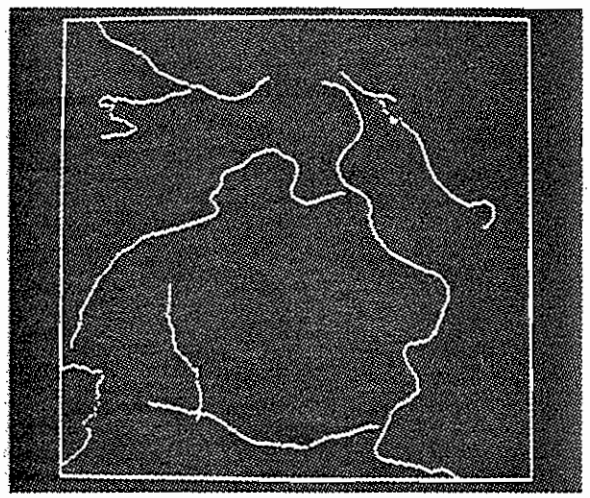
(c) THRESHOLDED LINEAR FEATURE POINTS



(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS

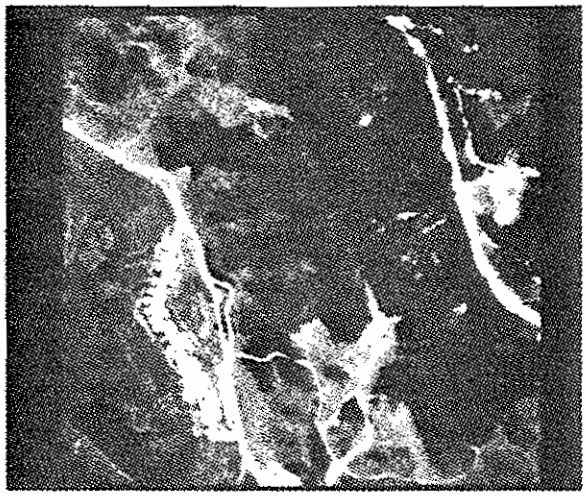


(e) SEGMENTS WITH 30 OR MORE POINTS

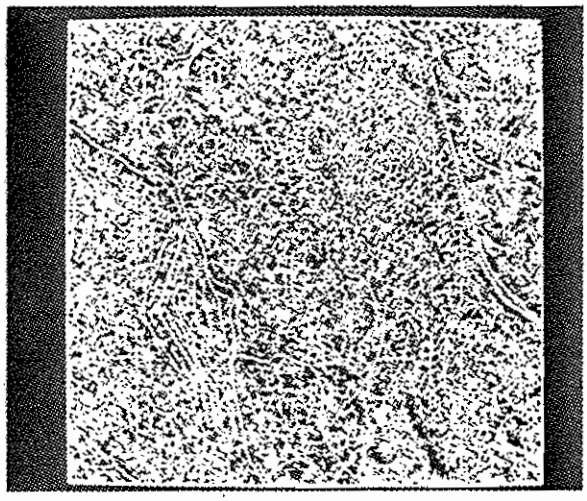


(f) 10 HIGHEST RANKED SEGMENTS

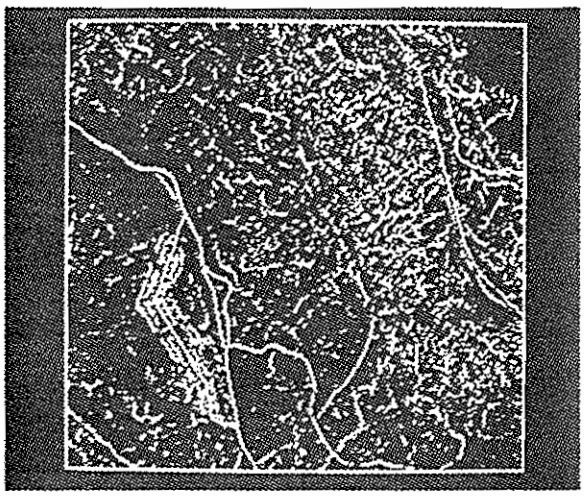
FIGURE 4 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A LOW RESOLUTION AERIAL IMAGE (DROAD2)



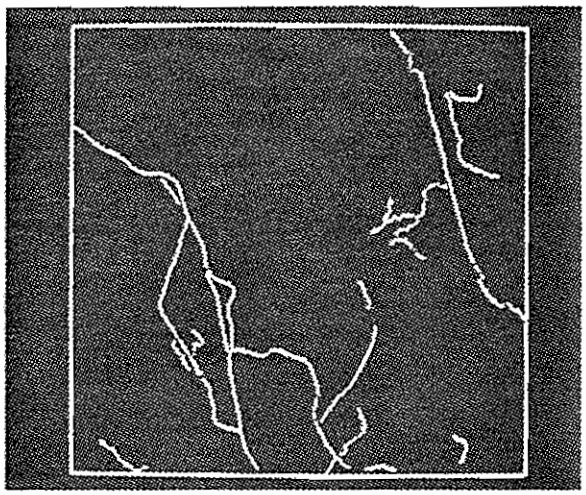
(a) ORIGINAL GRAY SCALE IMAGE: AFROAD1



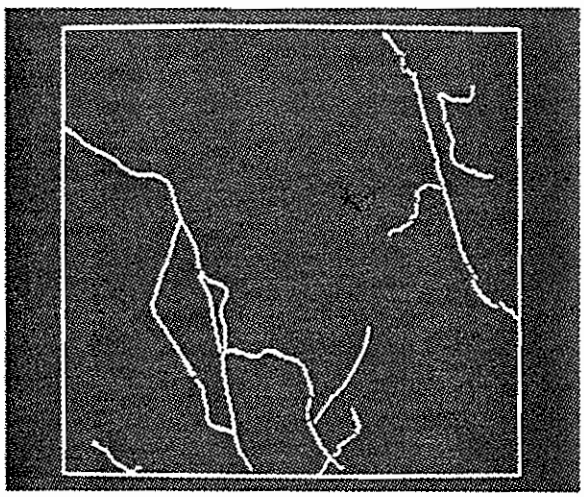
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS



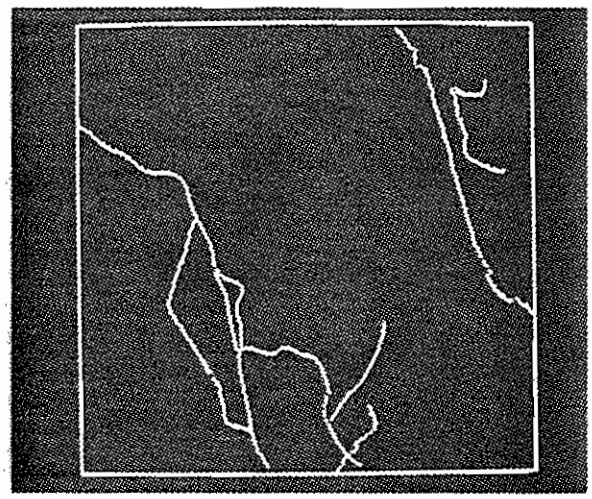
(c) THRESHOLDED LINEAR FEATURE POINTS



(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS

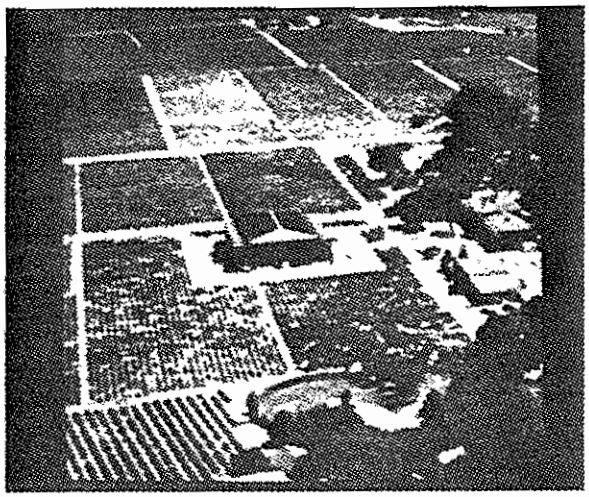


(e) SEGMENTS WITH 30 OR MORE POINTS

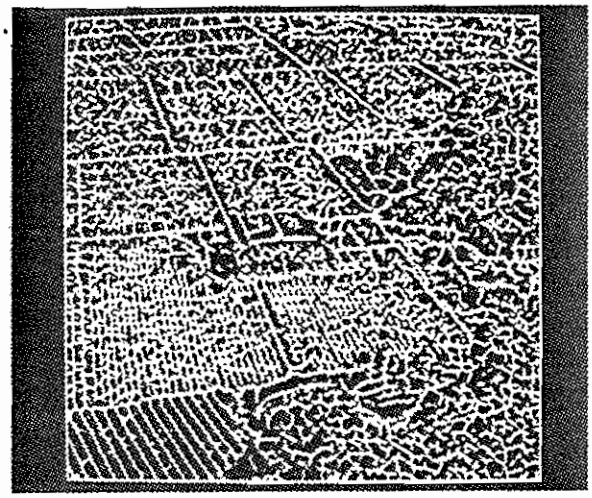


(f) 9 HIGHEST RANKED SEGMENTS

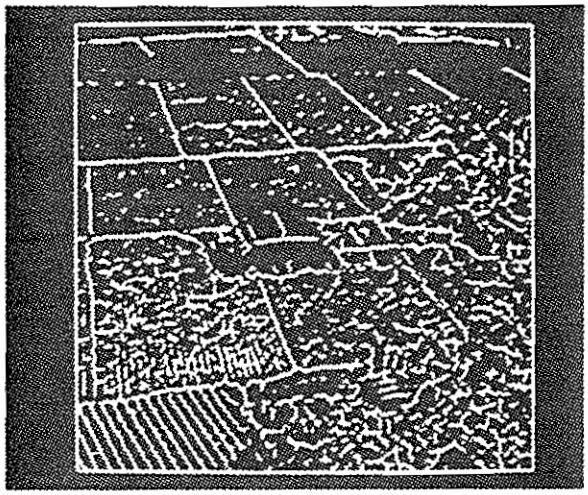
FIGURE 5 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A LOW RESOLUTION AERIAL IMAGE (AFROAD1)



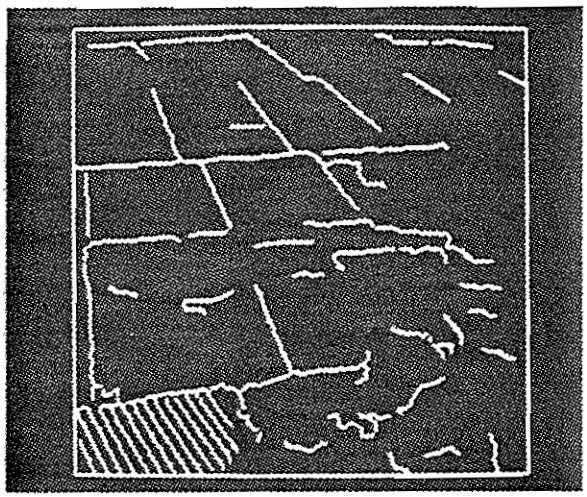
(a) ORIGINAL GRAY SCALE IMAGE: FARM-UL



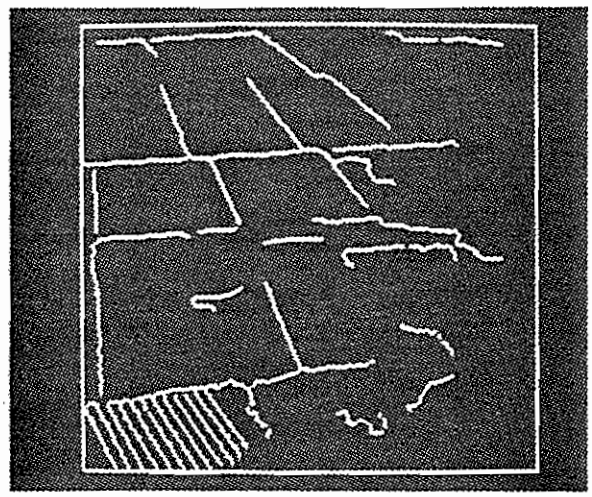
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS



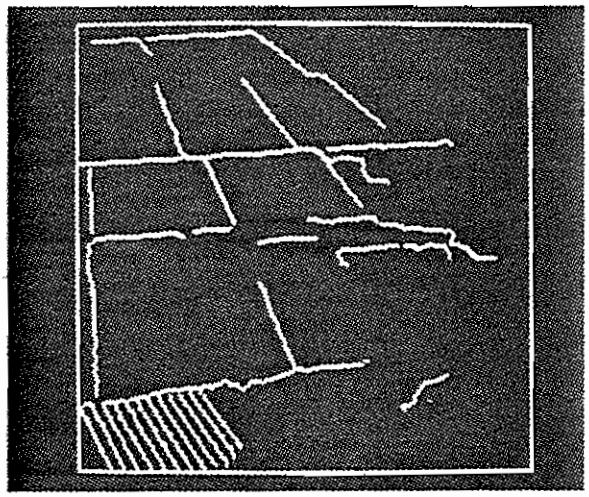
(c) THRESHOLDED LINEAR FEATURE POINTS



(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS



(e) SEGMENTS WITH 30 OR MORE POINTS



(f) 27 HIGHEST RANKED SEGMENTS

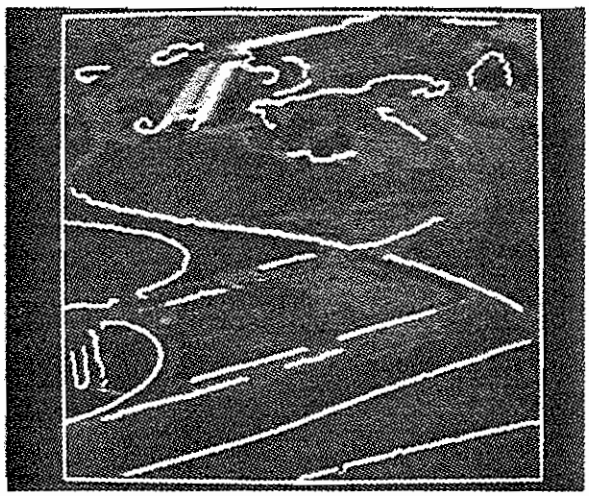
FIGURE 6 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A HIGH RESOLUTION AERIAL IMAGE (FARM-UL)

(a) ORIGINAL GRAY SCALE IMAGE: MARIN2

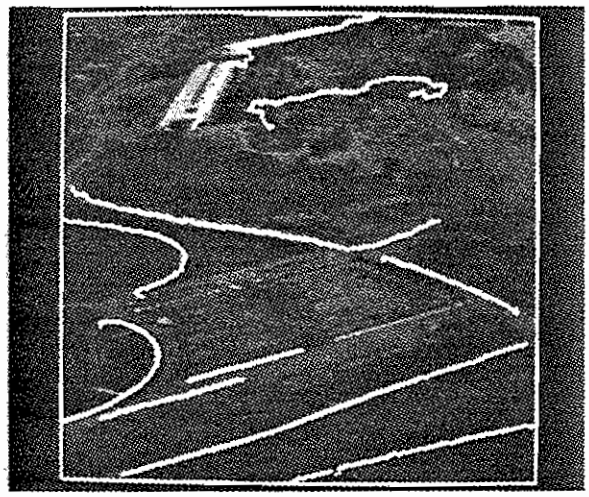
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS

(c) THRESHOLDED LINEAR FEATURE POINTS

(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS

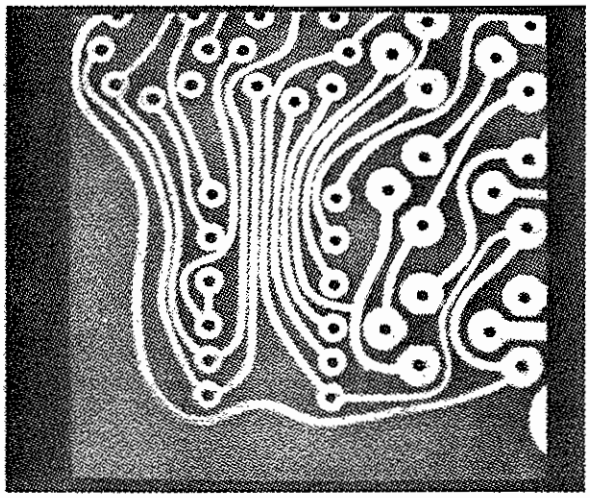


(e) SEGMENTS WITH 30 OR MORE POINTS

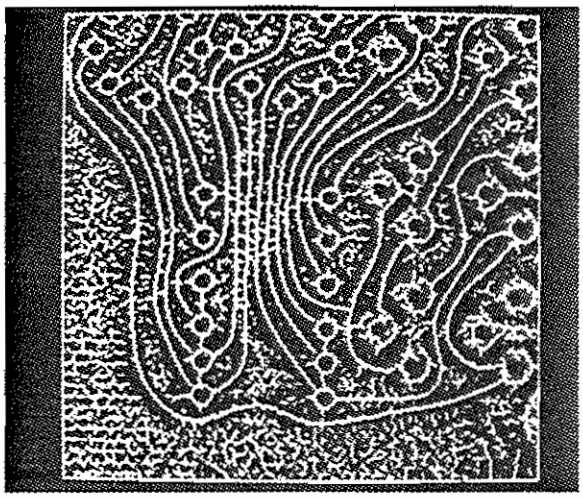


(f) 10 HIGHEST RANKED SEGMENTS

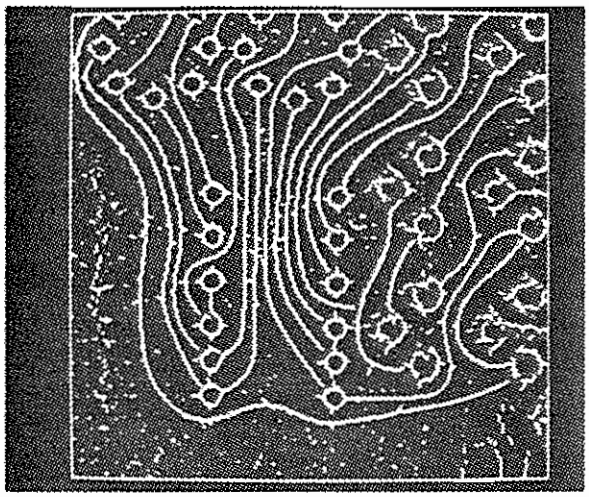
FIGURE 7 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A HIGH RESOLUTION AERIAL IMAGE (MARIN2)



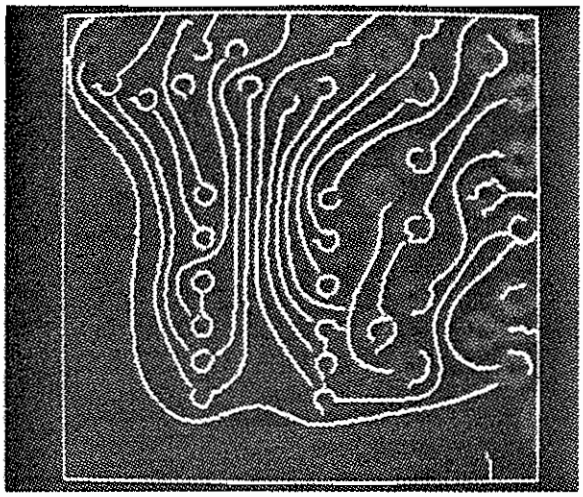
(a) ORIGINAL GRAY SCALE IMAGE: PCB1



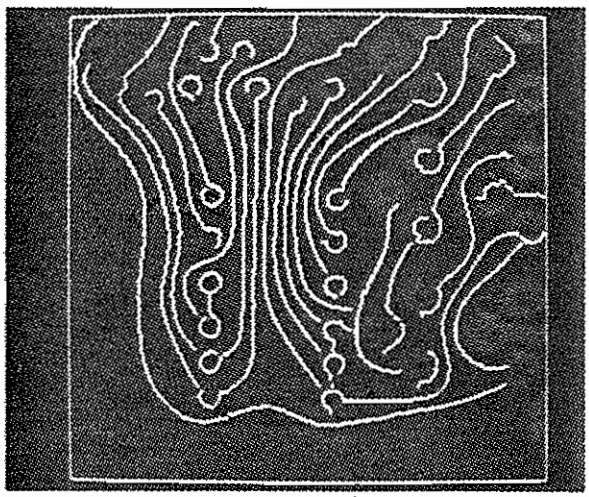
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS



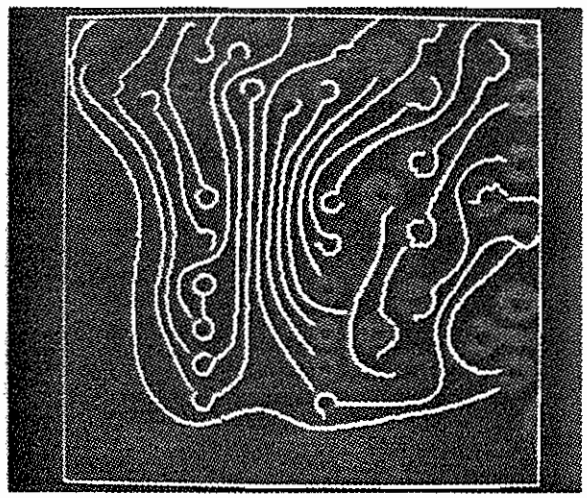
(c) THRESHOLDED LINEAR FEATURE POINTS



(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS

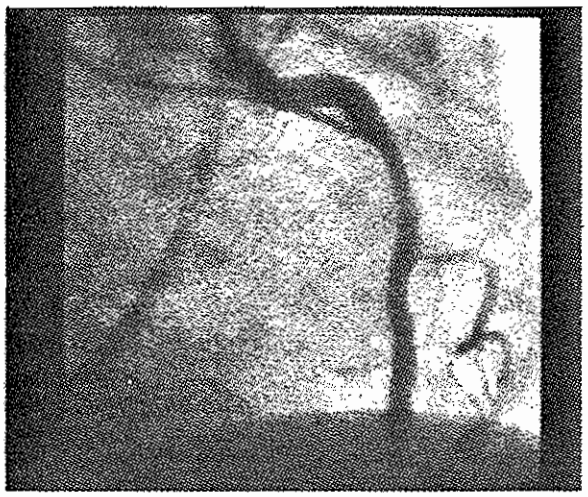


(e) SEGMENTS WITH 30 OR MORE POINTS

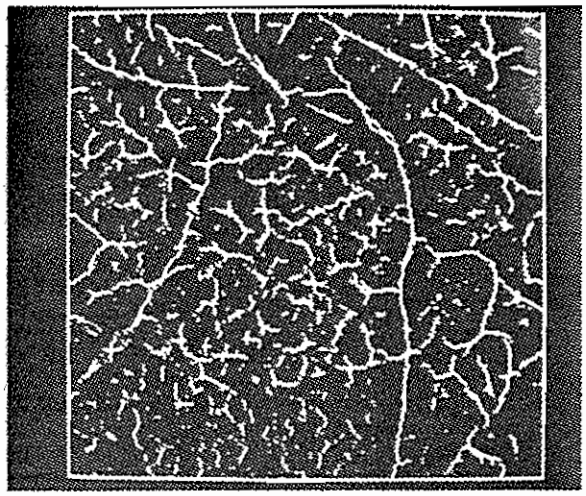


(f) 24 HIGHEST RANKED SEGMENTS

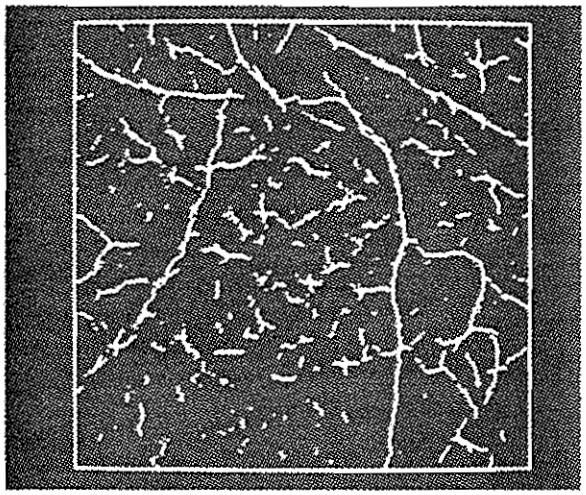
FIGURE 8 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A PRINTED CIRCUIT BOARD IMAGE (PCB1)



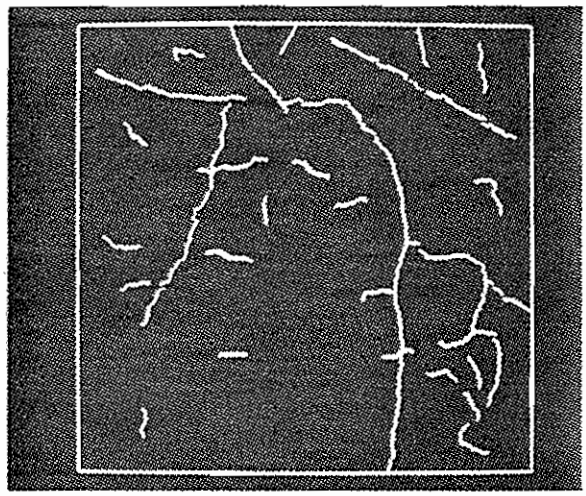
(a) ORIGINAL GRAY SCALE IMAGE: ANGIO5A



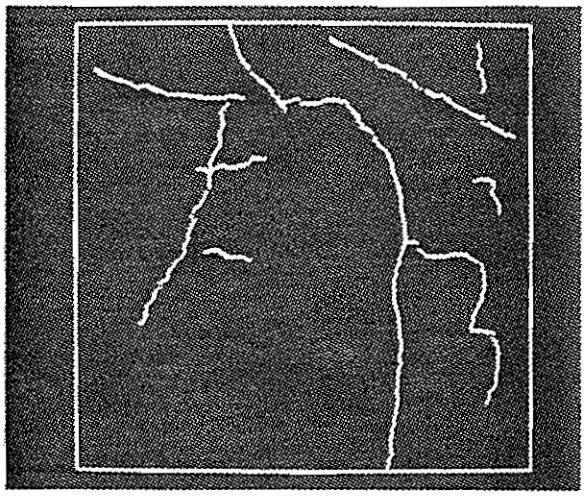
(b) DETECTED LINEAR FEATURE (INTENSITY MAXIMA) POINTS



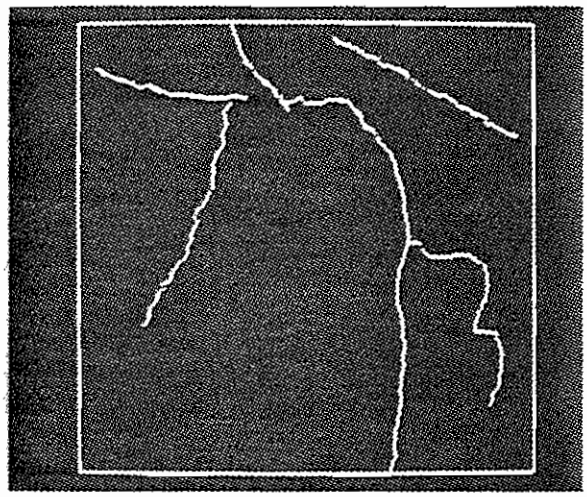
(c) THRESHOLDED LINEAR FEATURE POINTS



(d) SEGMENTS FOUND IN THRESHOLDED LINEAR FEATURE POINTS

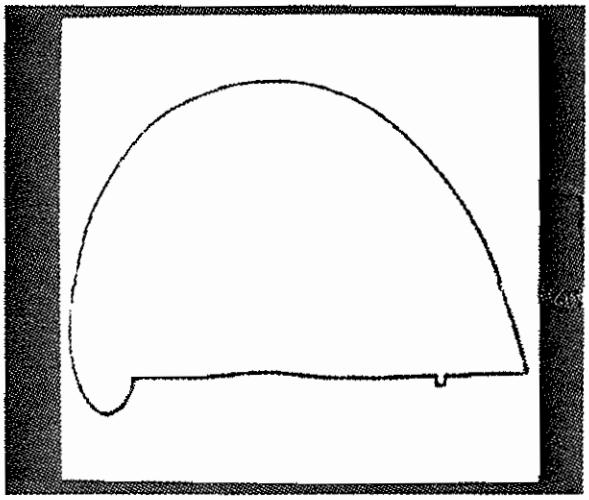


(e) SEGMENTS WITH 30 OR MORE POINTS

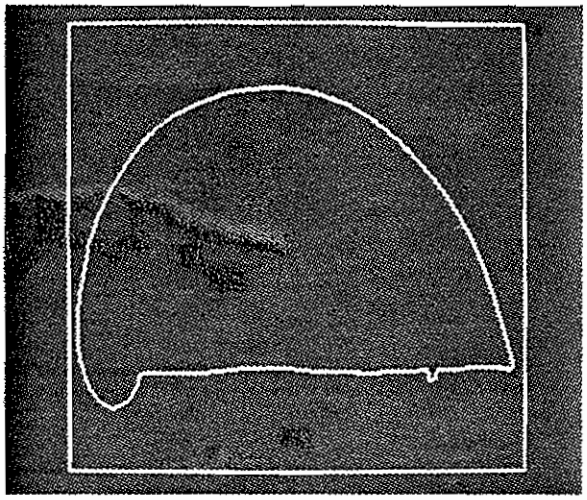


(f) 6 HIGHEST RANKED SEGMENTS

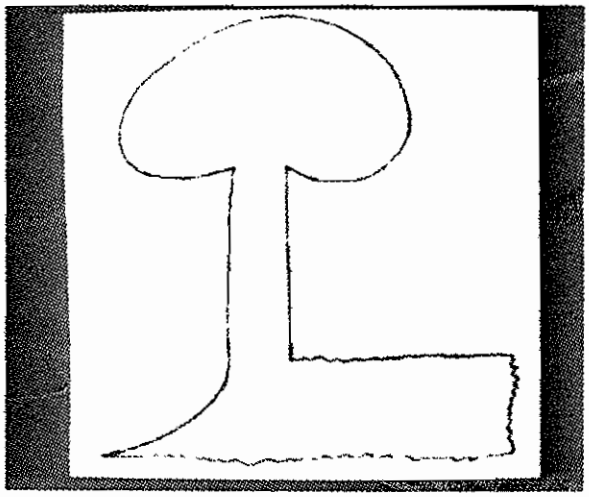
FIGURE 9 EXAMPLE SHOWING THE SEQUENCE OF STEPS IN THE DELINEATION PROCESS FOR A RADIOGRAPHIC ANGIOGRAM (ANGIO5A)



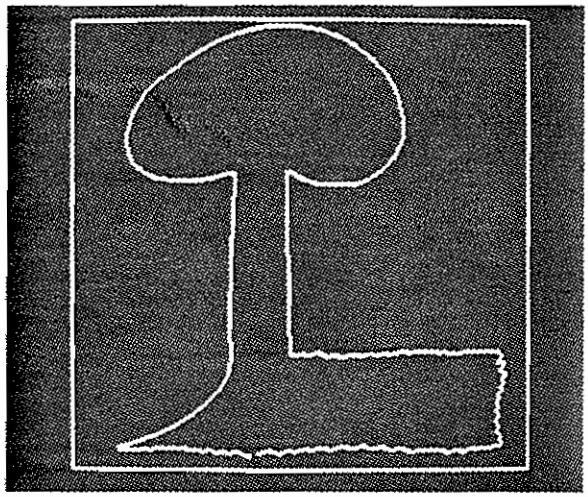
(a) ORIGINAL GRAY SCALE IMAGE: TURTLE



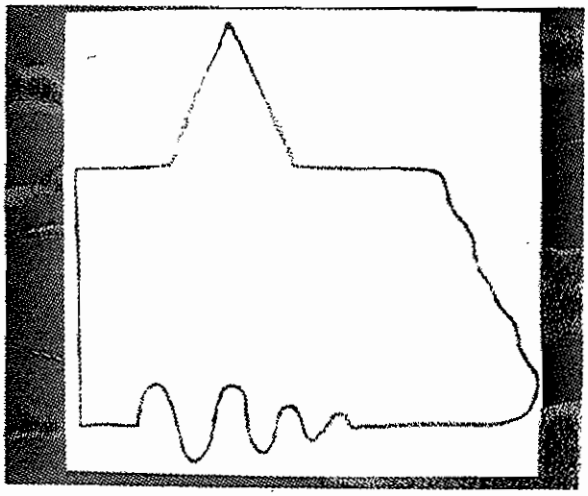
(b) EXTRACTED LINE SEGMENT FOR TURTLE



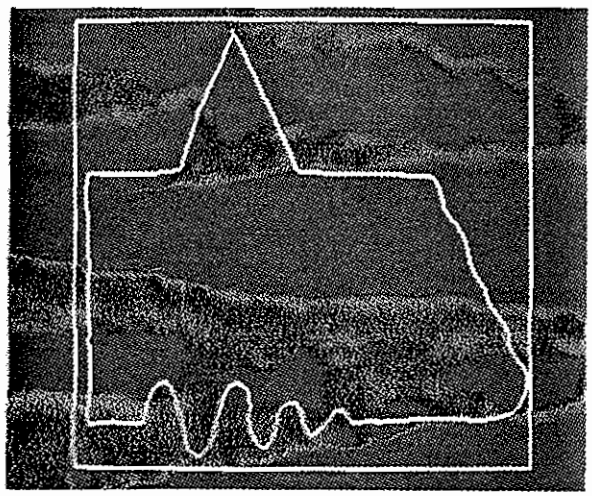
(c) ORIGINAL GRAY SCALE IMAGE: TREE



(d) EXTRACTED LINE SEGMENT FOR TREE



(e) ORIGINAL GRAY SCALE IMAGE: HOUSE



(f) EXTRACTED LINE SEGMENT FOR HOUSE

FIGURE 10 EXTRACTION OF LINEAR SEGMENTS FROM PENCIL DRAWINGS

```
    237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
169 143 144 146 143 144 146 146 149 147 150 153 158 151 155 153 154 154
170 147 149 145 145 148 148 146 142 146 148 148 149 151 153 149 151 154
171 212 214 211 205 213 206 209 207 199 203 200 202 197 195 192 193 192
172 230 229 228 226 228 227 228 228 230 230 227 223 228 223 219 221 224
173 233*236 229 231 231 229 230 233 229 232 228 225 231 230 223 223 225
174 224 228 226*233*232 232*231*235 232 229*231 231 231 233 223 224 226
175*234 229 229 229 229*238 231 233 229 231 225*232 228 229 230 230*234
176 227 229 225 229 231 231 229 230 229 227 224 230 232*234*237*236 230
177 225 224 227 232 225 228 228 229 226 227 226 225*234 233 234 231 230
178 231 223 228 227 227 227 229 226 228 229 230 226 227 230 233 230 229
179 223 229*235 232 229 231 230 229 227 231 231 231 227 229 224 230 230
180 226 226 231 228 231 229 229 226*234*233 228 231 224 226 230 228 226
181 183 186 186 187 191 192 191 193 197 198 198 198 196 197 195 203 202
182 153 153 149 149 151 151 153 158 161 161 164 158 161 163 168 169 173
183 143 142 141 140 141 143 144 147 148 147 150 155 150 159 159 154 158
```

a. Gray scale intensities of a thick, noisy, horizontal line in window 237,169 253,183 of unsmoothed image, PCB1.
Line has intensity values greater than 200.
*-ed pixels indicate points at which global maxima is greater than 15

```
    237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
169 164 168 172 175 177 178 179 179 179 180 180 179 177 174 170 167 166
170 176 177 179 180 181 181 181 181 181 181 181 181 180 178 176 175 175
171 195 195 195 195 195 195 195 195 194 194 194 193 193 191 190 190 190
172 212 212 212 211 211 211 211 211 211 210 210 209 208 207 206 206 206
173 223 223 223 223 223 223 223 223 223 222 221 221 220 219 218 218 218
174 227 227 228 228 228 229 229 229 228 228 227 227 227 226 226 226 226
175*228*228*229*229*230*230*230*230*229*229 228 229 229 230 229 229 229
176 227 228 228 229 229 230 230 230 229 229*228*229*230*231*231*231*231
177 226 227 228 228 228 229 229 229 228 228 228 229 229 230 230 231 231
178 224 225 225 226 226 226 227 227 226 226 226 227 227 228 228 229 228
179 217 218 219 220 220 220 221 221 221 221 221 221 222 222 222 223 223
180 205 206 206 207 207 208 208 209 210 210 211 210 210 211 211 212 213
181 187 188 189 190 190 191 192 193 194 195 195 195 195 195 196 197 199
182 171 172 174 175 176 176 178 178 179 180 181 181 180 181 181 182 183
183 163 166 169 171 172 173 174 174 175 175 175 175 174 173 172 171 172
```

b. Gray scale intensities after smoothing with mask of diameter 11.
*-ed pixels indicate points at which global maxima is greater than 15

FIGURE 11 SMOOTHING WITH AN APPROPRIATE MASK WILL ENHANCE LINE STRUCTURE

| | 167 | 168 | 169 | 170 | 171 | 172 | 173 | 174 | 175 | 176 | 177 | 178 | 179 | 180 | 181 | 182 | 183 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 43 | 85 | 92 | 104 | 112 | 122 | 100 | 83 | 79 | 81 | 75 | 75 | 70 | 76 | 77 | 69 | 64 | 75 |
| 44 | 146 | 121 | 105 | 107 | 122 | 111 | 100 | 83 | 83 | 86 | 79 | 72 | 73 | 78 | 68 | 67 | 67 |
| 45 | *153 | *165 | *170 | 161 | 152 | 141 | 136 | 140 | 136 | 134 | 128 | 114 | 89 | 71 | 70 | 73 | 66 |
| 46 | 124 | 125 | 164 | *168 | *173 | *174 | *170 | *170 | *166 | *173 | *182 | *181 | *162 | 122 | 72 | 65 | 67 |
| 47 | 111 | 130 | 135 | 143 | 142 | 144 | 143 | 143 | 141 | 139 | 130 | 132 | *154 | *148 | 119 | 86 | 69 |
| 48 | 94 | 106 | 130 | 135 | 142 | *146 | 141 | 144 | 143 | 142 | 131 | 123 | 126 | 117 | *130 | *137 | 98 |
| 49 | 93 | 98 | 109 | 137 | 145 | 147 | 154 | *159 | 153 | 153 | 143 | 141 | 135 | 134 | 109 | *138 | 127 |
| 50 | 96 | 103 | 112 | 121 | 142 | 143 | 142 | 144 | *156 | 154 | 142 | 144 | 144 | 144 | 132 | 123 | *141 |
| 51 | 109 | 117 | 115 | 122 | 128 | 146 | 142 | 148 | 146 | *153 | 146 | 142 | 143 | 141 | 138 | 133 | 123 |
| 52 | 110 | 124 | 124 | 126 | 133 | 133 | 142 | 147 | 147 | *151 | 146 | 147 | 146 | 140 | 129 | 129 | 117 |
| 53 | 109 | 117 | 121 | 117 | 129 | 130 | 144 | 152 | 147 | 147 | 148 | 150 | 145 | 144 | 138 | 133 | 123 |
| 54 | 103 | 109 | 111 | 107 | 119 | 134 | 136 | 144 | 143 | *146 | 144 | 140 | 141 | 146 | 139 | 139 | 126 |
| 55 | 101 | 105 | 113 | 127 | 139 | 139 | 129 | 139 | 130 | 142 | *149 | 146 | 139 | 136 | 139 | 146 | 134 |
| 56 | 110 | 111 | 127 | 130 | 127 | 136 | 129 | 131 | 127 | 134 | 141 | *147 | 144 | 147 | 127 | 137 | 140 |
| 57 | 124 | 127 | 121 | 125 | 123 | 126 | 127 | 128 | 127 | 131 | 139 | 136 | 148 | 143 | 137 | 133 | 133 |

a. Gray scale intensities of parts of two segments in window 167,43 183,57 of unsmoothed image, DROAD2.
*-ed pixels indicate points on two distinct linear segments.

| | 167 | 168 | 169 | 170 | 171 | 172 | 173 | 174 | 175 | 176 | 177 | 178 | 179 | 180 | 181 | 182 | 183 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 43 | 110 | 112 | 113 | *114 | 113 | 111 | 107 | 103 | 100 | 97 | 94 | 91 | 87 | 83 | 79 | 76 | 74 |
| 44 | 115 | 118 | 120 | *122 | 121 | 120 | 117 | 113 | 110 | 107 | 103 | 99 | 94 | 89 | 84 | 80 | 76 |
| 45 | 118 | 122 | 126 | 128 | *129 | 129 | 127 | 124 | 121 | 118 | 114 | 109 | 103 | 97 | 90 | 85 | 80 |
| 46 | 118 | 124 | 129 | 133 | 135 | *136 | 135 | 133 | 131 | 128 | 125 | 119 | 113 | 106 | 98 | 92 | 86 |
| 47 | 117 | 123 | 129 | 134 | 138 | 140 | *140 | 140 | 139 | 136 | 133 | 128 | 122 | 115 | 107 | 99 | 92 |
| 48 | 114 | 120 | 126 | 133 | 138 | 141 | *143 | *143 | 143 | 141 | 138 | 134 | 128 | 122 | 115 | 107 | 99 |
| 49 | 113 | 118 | 124 | 130 | 136 | 140 | 143 | *144 | *145 | 143 | 141 | 138 | 133 | 128 | 121 | 114 | 106 |
| 50 | 112 | 116 | 122 | 128 | 134 | 139 | 142 | 144 | *145 | 144 | 143 | 140 | 137 | 132 | 126 | 120 | 111 |
| 51 | 113 | 116 | 121 | 126 | 132 | 137 | 140 | 143 | *145 | *145 | 144 | 142 | 139 | 135 | 130 | 124 | 116 |
| 52 | 114 | 116 | 120 | 125 | 130 | 135 | 139 | 142 | 143 | *144 | 144 | 142 | 140 | 137 | 132 | 127 | 119 |
| 53 | 115 | 117 | 120 | 124 | 129 | 133 | 137 | 140 | 142 | *143 | 143 | 142 | 141 | 138 | 134 | 129 | 121 |
| 54 | 118 | 118 | 120 | 123 | 128 | 131 | 135 | 137 | 140 | 141 | *142 | 142 | 141 | 139 | 135 | 131 | 124 |
| 55 | 121 | 120 | 121 | 123 | 126 | 130 | 132 | 135 | 137 | 139 | 140 | *140 | 140 | 139 | 136 | 132 | 125 |
| 56 | 124 | 122 | 122 | 123 | 125 | 127 | 130 | 132 | 134 | 136 | 138 | 139 | *139 | 139 | 137 | 133 | 127 |
| 57 | 128 | 124 | 122 | 123 | 124 | 125 | 127 | 128 | 131 | 133 | 136 | 137 | *138 | 138 | 137 | 134 | 129 |

b. Gray scale intensities after smoothing with mask of diameter 21.
*-ed pixels indicate points on the linear segment found.
Upper right segment has been eliminated.

FIGURE 12 SMOOTHING WITH A LARGE DIAMETER MASK MAY ELIMINATE OR SHIFT SEGMENTS